BBC RD 1976/31

BBC

RESEARCH DEPARTMENT

1974
THE QUEEN'S AWARD TO INDUSTRY

REPORT

60/120 Mbit/s multiplexing equipment for high-quality video and audio signals

J.H. Stott, M.A.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION November 1976

BBC RD 1976/31
UDC 621.396.41:
621.396.132

60/120 MBIT/S MULTIPLEXING EQUIPMENT FOR HIGH-QUALITY VIDEO AND AUDIO SIGNALS

J.H. Stott, M.A.

Summary

If digital transmission is applied to television programme distribution it is desirable to be able to convey both picture and sound information in a single digital circuit. Multiplexing of digitally coded video and sound signals is therefore necessary. A technique known as justification is required when such digital signals with independent bit rates are multiplexed together.

Following a brief review of justification techniques a description is given of the design of equipment which was initially produced for use in field trials with a UK Post Office experimental 120 Mbit/s digital line transmission system. The equipment combines two 60 Mbit/s packages for transmission at 120 Mbit/s. Each 60 Mbit/s package can carry one System I PAL colour video signal and six high-quality sound signals.

Issued under the authority of

G.D. Montwater

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

November 1976

(EL-122)

60/120 MBIT/S MULTIPLEXING EQUIPMENT FOR HIGH-QUALITY VIDEO AND AUDIO SIGNALS

60/120 MBIT/S MULTIPLEXING EQUIPMENT FOR HIGH-QUALITY VIDEO AND AUDIO SIGNALS

J.H. Stott, M.A.

1. Introduction

This Report describes the design of equipment which assembles two similar 60 Mbit/s data streams, each providing one video channel and six high-quality sound channels, and combines them to provide a single 120 Mbit/s data stream.

The equipment was primarily intended for use with an experimental 120 Mbit/s digital line system installed for the UK Post Office (UK PO) between Guildford and Southampton via Portsmouth, where, by courtesy of the UK PO, the BBC carried out field trials in 1975.[1] However, the possibility of trials with other transmission systems, e.g. a satellite link, was foreseen and allowed for during the evolution of the equipment.

The main aim of this Report is to describe the general requirements of the multiplexer and the design approach used to meet them.

2. Initial design considerations

2.1. Preliminary history

The initial incentive to produce multiplexing equipment was given by the prospect of access to the UK PO experimental digital line transmission system for joint BBC/PO tests in 1975.

As the capacity of the line system was 120 Mbit/s it was clearly possible to convey one video signal using well-known techniques, namely 8-bit linear pulse-code modulation (p.c.m.) with around 3 x f_{sc}* sampling frequency. When one parity bit per sample is added to provide error protection using a concealment technique[2] the bit rate required is in the region of 120 Mbit/s. Such a system was in fact built by BBC Research Department for demonstration of transmission by circular waveguide in September 1970 at the Post Office Research Department, Martlesham,[3] but a new version of the system was built for use in the April 1975 BBC/PO pilot tests at Portsmouth.[1]

The possibility of gaining further experience from the line system experiments was offered by two contemporary developments:

(i) the invention of a sub-Nyquist, 2 f_{sc} sampling frequency technique for coding a composite PAL video signal;[4]

and (ii) the development of 2048 kbit/s digital sound multiplex equipment.[5]

Devereux[6] demonstrated that existing differential p.c.m. (d.p.c.m.) and hybrid d.p.c.m.* (h.d.p.c.m.) video coding techniques could be used in conjunction with sub-Nyquist sampling to offer a substantial total saving in bit rate. For example, the combination of 2 f_{sc} sampling with 6-bit d.p.c.m. or h.d.p.c.m. coding requires only 53·2 Mbit/s, although some additional bit rate is desirable for error correction coding.[7,8] Two such video signals may be accommodated within 120 Mbit/s with capacity to spare; sending extra signals thus becomes possible.

Any digital transmission system eventually adopted for television programme distribution will carry both video and audio signals; experience of multiplexing such signals in digital form was therefore desirable. The 2048 kbit/s sound multiplex equipment appeared a convenient source of digital sound signals for the experiment since it was already under development by BBC Designs Department.

It appeared attractive to form a 120 Mbit/s signal from two similar 60 Mbit/s 'packages', each including one video channel and making some provision for high-quality sound. This idea was supported by the prospect of further trials with other transmission systems, some of which would be restricted to about 60 Mbit/s. One example is digital transmission through a satellite transponder with 40 MHz nominal bandwidth.

The general objectives thus become clear:

(i) to produce apparatus for use in the line system experiments which conveyed both video and audio signals, thus gaining experience in multiplexing;

and (ii) to form a 120 Mbit/s stream from two similar 60 Mbit/s streams, each containing video and audio signals. This would cover any future experiments at either 60 or 120 Mbit/s, and would also give the experience of transmitting reduced bit rate video signals over a real transmission system.

Fig. 1 is a block diagram of the sending equipment.

2.2. Details

The inputs and outputs of the multiplexing equipment will now be specified in greater detail.

The UK PO digital line system uses the HDB3 ternary code[9] for interfacing at 120 Mbit/s; separate interfacing units were constructed to convert between this and binary NRZ data with clock. A tolerance of ±15 p.p.m. (parts per million) is specified for the clock frequency, nominally 120 MHz.

* f_{sc} = colour subcarrier frequency = 4·43361875 MHz for the PAL 625-line systems.

* In this context, 'hybrid d.p.c.m.' means that both p.c.m. and d.p.c.m. code-words are transmitted.[6]

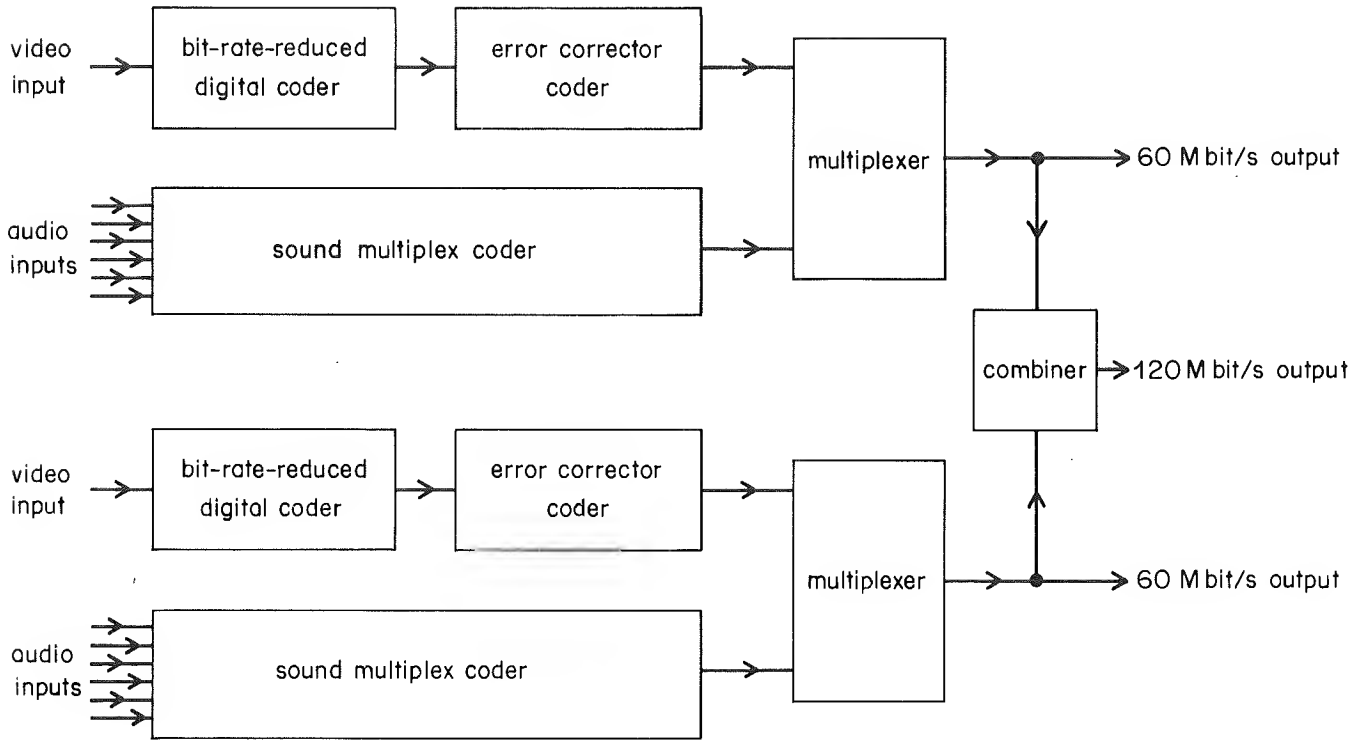


Fig. 1 - Sending equipment: general block diagram

The sound multiplex equipment produces a serial stream at 2048 kbit/s ±50 p.p.m. The six sound signals which it may carry are each sampled at 32 kHz and quantised to 13-bit accuracy; the samples are then compressed to $10^1/16$ bits/sample using near-instantaneous companding.[5] The equipment incorporates its own provision for error protection using a concealment technique.

The video coding equipment samples the video signal at twice the colour subcarrier frequency and quantises each sample to 8-bit accuracy. The 8-bit samples are then reduced to 6 bits using d.p.c.m. or h.d.p.c.m. coding. The digital input to the transmission equipment is thus in parallel form, comprising 6-bit words at $2f_{sc}$. The tolerance of the clock frequency is therefore that of the colour subcarrier, namely ±0·25 p.p.m. (national standard) or ±1·1 p.p.m. (international standard).

An error corrector coder increases the video bit rate before the multiplexer proper. The extent to which the bit rate is so increased forms part of the process of system design and is considered in Section 3.

The input and output bit rates of the multiplexer discussed here are unrelated and indeed may vary within their individual tolerances independently of each other. To multiplex such asynchronous signals requires the use of a technique known as justification. This will be described in Section 4.

3. Choice of video coding and error correction coding

Clearly, the sum of the bit rates of the individual signals which are to be multiplexed to produce one 60 Mbit/s signal must not exceed 60 Mbit/s. If we allocate 2048 kbit/s for a sound-signal multiplex 57·952 Mbit/s remain for the video signal (including error-correction coding) and any necessary control signals for framing and justification.

Step 1 What are the possible basic video bit rates?

Assuming $2f_{sc}$ (i.e. sub-Nyquist) sampling and some form of differential coding giving an integer m bits/sample we obtain the following table:

m	bit rate (Mbit/s)
4	35·47
5	44·34
6	53·2
7	62·07

As m is reduced either the picture quality suffers or the cost of a codec (coder and decoder) increases, or both. Thus the highest usable value of m is to be preferred.

6 bits/sample, occupying 53·2 Mbit/s, seemed a suitable choice and was attractive as a suitable codec was already available. Operation at a reduced number of bits per sample is of course possible for comparison in which case dummy information would be sent in the unused bit positions. In fact, an h.d.p.c.m. option using 5 bits/sample was included in the equipment.

Step 2 What form of error-correction coding should be employed?

Previous work[7,8] demonstrated the effectiveness of the Wyner-Ash (8, 7) code in correcting errors in a digital

video d.p.c.m. or h.d.p.c.m. signal. Using this form of error-correction coding on all six bits in each sample, the total video bit rate becomes:

(53·2 x 8)/7 = 60·8 Mbit/s — this is too great.

To keep within the desired bit rate the following courses are possible:

(i) use the Wyner-Ash (8, 7) code, but applied to the more significant bits only (as in previous experiments[7,8]);

(ii) apply correction to all 6 bits, but use the next less redundant Wyner-Ash code — the (16, 15) code;

or (iii) adopt a completely new error-correcting strategy.

The previous investigation had selected the Wyner-Ash code as a suitable compromise between performance and complexity. This conclusion was still considered to be valid, moreover practical experience has been gained in the use of the code. Course (iii) was therefore rejected.

The previous results[8] suggested that using the (8, 7) code to correct all 6 bits of a d.p.c.m. or h.d.p.c.m. word, rather than the four most significant bits only, as done in the earlier work, would offer a worthwhile improvement in the performance. Furthermore this improvement would be sufficiently large that it would not be nullified by a change to the slightly less effective* (16, 15) code. In other words, 6-bit correction using the (16, 15) code offers a better performance than correction of only the four most significant bits using the (8, 7) code.

A (16, 15) codec is slightly more complex than one for the (8, 7) code, but this is balanced by the awkward arrangement of data and parity bits which would be entailed with correction of only 4 bits out of 6 with the (8, 7) code.

Using the (16, 15) code in our application makes the total video bit rate equal to:

(53·2 x 16)/15 = 56·75 Mbit/s.

This represents a comfortable fit within the 57·95 Mbit/s available for video and control signalling. With the above discussion taken into consideration the (16, 15) code was chosen.

4. Justification

Before proceeding further with the system design a brief description of 'justification', sometimes known as 'pulse stuffing', will be given.

4.1. The need for justification

Suppose that bit stream A ('the input') of bit-rate f_A is to be incorporated into another stream B ('the output') of bit rate f_B ($f_B > f_A$). Clearly, if f_A bears a fixed rational relationship to f_B,

i.e. if $f_A : f_B = p : q$ where p and q are integers,

then a cyclic format for stream B could be contrived. For example, every q bits in stream B could contain p bits from stream A and $(q - p)$ other bits which could either be derived from other input channels or merely fill the space with dummy information. We will see later (Section 5.5) that the problem of multiplexing data and parity information in the error-corrector coder is of this form.

However, where no fixed rational relationship between f_A and f_B holds, such a fixed cyclic format is not possible. In such a situation justification is required.

4.2. General principles of justification

The basic idea of justification techniques is to allocate a dual function to certain digit time slots* in the output stream. These may carry either input information or dummy information *on a time-varying basis* as appropriate to ensure that the average rate of transmission of input information equals the input bit rate. Obviously some form of signalling must also be provided to indicate whether these time slots carry real or dummy information at any given time.

Such dual function time slots are known as *'justifiable time slots (JTS)'* and the signalling as to their content is carried by a *'justification control signal (JCS)'*.

The output stream is usually divided into frames of fixed length within which certain digit time slots are designated as JTS or for the JCS. If there is more than one input, as in multiplexing, then each input has its own JTS and JCS positions allocated within the frame. With such a frame structure, two implementations are often used, as discussed in the literature:[10] 'positive justification' and 'positive/negative justification'. Other systems are possible, however.**

4.3. Positive justification

For this system each input is allocated one JTS per frame. Clearly, the maximum permissible input bit rate is transmitted when input data occupies the JTS in every frame; conversely, the minimum input bit rate will require sending dummy information in the JTS in every frame. The range of operation must always lie between these extremes if no data is to be lost, and normally the frame length is chosen so that neither extreme is approached very closely.

* Using a (16, 15) code in place of (8, 7) reduces the acceptable transmission path error rate (for the same decoded output error rate) by a factor of about 0·62 (see Reference 7).

* A digit time slot is a cyclic time interval, which can be recognised and defined uniquely, allocated to a single digit (CCITT Rec.702).

** For example, a system devised by J.P. Chambers was used for pilot field trials[1] in which 119·7 Mbit/s of information were to be carried in 120 Mbit/s. In that system it was the *frame length* which was changed as necessary on a time-varying basis.

We may define a few terms, paraphrasing CCITT Recommendation 702:

The *justification rate* is the average rate at which dummy information is carried in the JTS.

The *justification ratio* is the ratio of the justification rate to the maximum possible justification rate (when every JTS is dummy). In other words, the justification ratio represents the fraction of the time that the JTS carries dummy information.

The *nominal justification rate* and the *nominal justification ratio* are the values of the above parameters when the input and output bit rates assume their nominal values.

4.4. Positive/negative justification

For this technique each input is given two JTS locations in every frame, JTS 1 and JTS 2.

When input and output rates have their nominal values, one JTS, say JTS 1, carries input information and the other, JTS 2, carries control information. When the input rate is higher than nominal (or, equally, the output rate is lower than nominal) JTS 2 will sometimes also carry input information. Conversely, when the input rate is lower than nominal (or, equally, the output rate is higher than nominal) JTS 1 will sometimes carry dummy information.

The nominal justification rate is thus effectively zero; as input and output rates vary over their tolerances both positive and negative values of justification rate will occur.

Note that if nominal input and output rates were always maintained, the system simplifies to the form first discussed in Section 4.1 above. The purpose of justification applied in this way would be merely to absorb the tolerances of the input and output rates whose nominal values must have a rational relationship. If no rational relationship exists, then positive justification would be required even if input and output rates remained exactly constant at their nominal values.*

* Positive/negative justification could be used with non-integer relationships; however, there seems little reason to do so. The system would amount to positive justification for nominal rates, and the region around nominal, and would only have a true positive/negative function when at one extreme of the range.

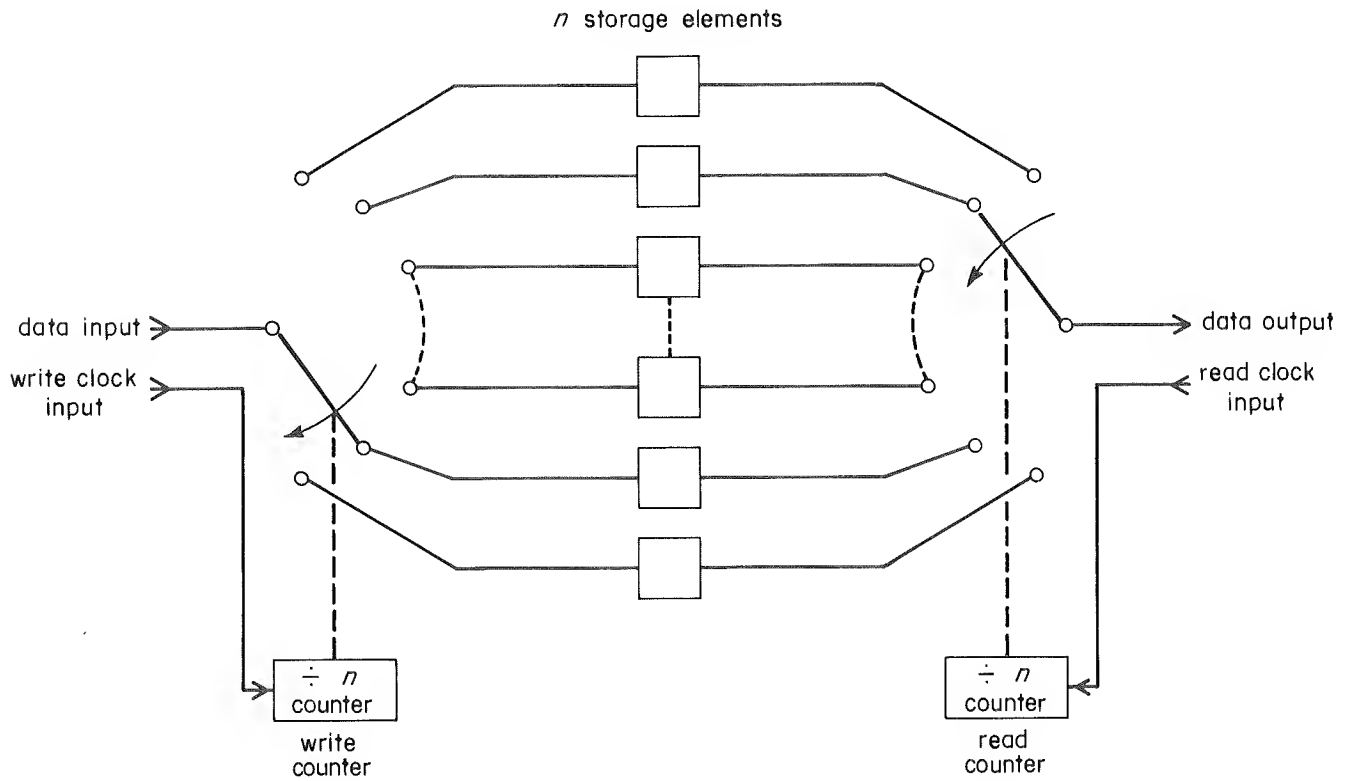


Fig. 2 - n-stage elastic store arrangement

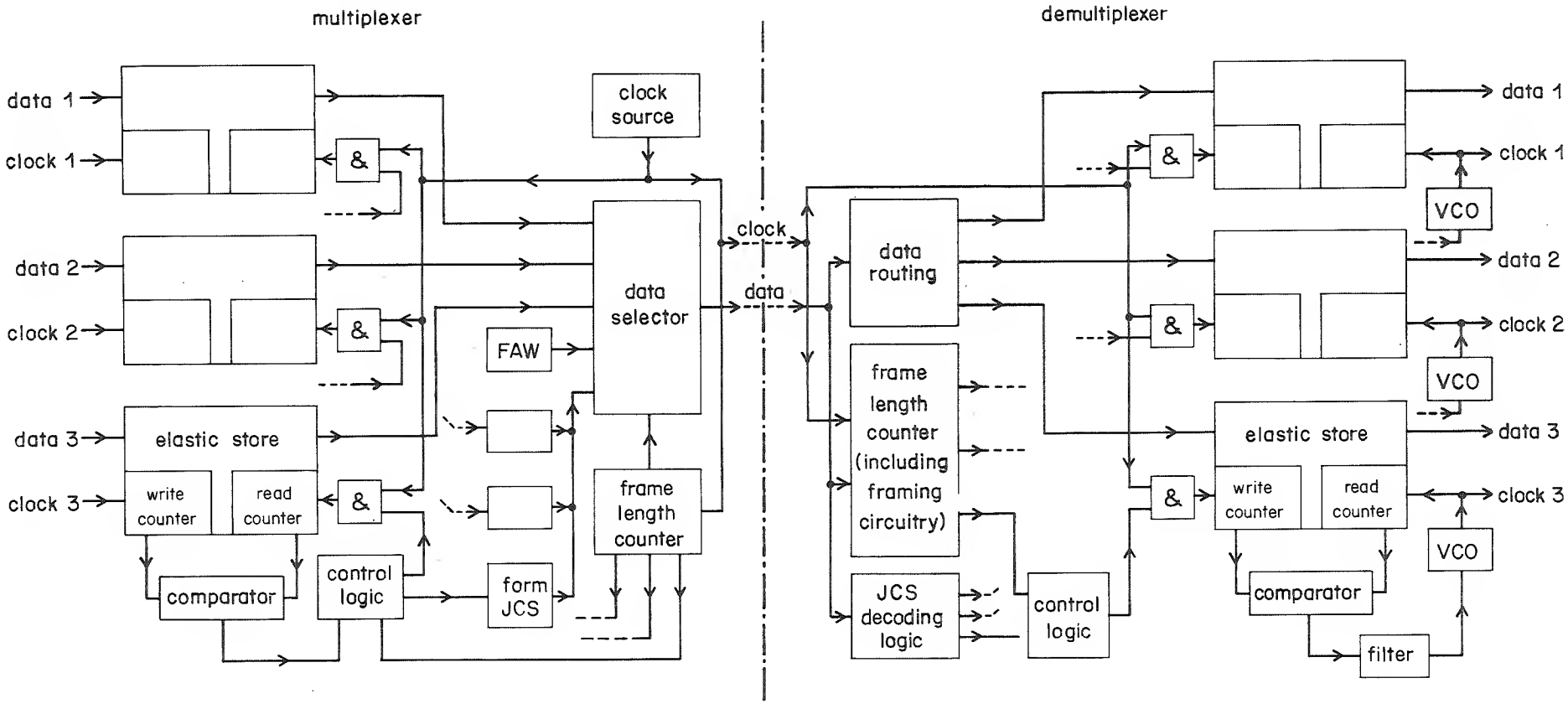


Fig. 3 - Illustration of multiplexing using justification

Positive/negative justification would appear to have greatest application in the situation where a designer can choose related input and output nominal bit rates. A possible example is the design of a telephone/telecommunication system where a multiplex hierarchy is built up from the lower bit rates upwards. The number of input channels to be multiplexed and the nominal output rate at each level of the hierarchy are parameters over which the designer has some control.

Positive/negative justification appears rather more complicated to instrument than positive justification, seemingly without offering any useful advantages.*

4.5. Apparatus

Having outlined the general principles of justification, we now consider how to apply them in practice, where input and output nominal bit rates are unrelated, and where positive justification is therefore used.

The most important point to note is that we are considering a situation where input information is provided at a steady rate but is transmitted in an irregular fashion; for any given input the average rate of information transmission equals the input rate but the instantaneous rate varies. This implies a need for storage. An approximate analogy is a gas-holder. This is steadily provided with gas from the gas-producing plant but is subjected to a fluctuating demand. The average consumption must equal the rate of production or the gas-holder will become completely full or empty. Furthermore the capacity of the gas-holder must be large enough to cope with the prevailing short-term changes of demand (e.g. the cooking of Sunday lunch). We may think of this as corresponding to the gaps in frame structure where a frame-alignment sequence, justification-control signal or information from another input is transmitted. The gas-holder cannot cope with a long-term change of demand (e.g. from summer to winter) and in this case the rate of production would be changed. The analogy is slightly imperfect at this point since a change in the ratio of input and output rates is accommodated by a change of justification ratio — in other words an increase of input rate is matched by an increase in average transmitted rate.

Thus we require storage of variable capacity — sometimes called 'elastic' storage. Fig. 2 shows a typical arrangement for providing this. Data is written into n storage elements in turn, directed by a $\div n$ counter, the 'write counter'. Another counter, the 'read counter', causes data to be read out of the stores in turn at the rate demanded by the read clock. If no data is to be repeated or lost the switches must not overtake each other. The amount of data in the store at any time is given by the number of switch positions by which the write counter leads the read counter. As with the gas-holder, the store must lie between empty and full without quite reaching the extremes.

Fig. 3 shows how elastic stores are incorporated into a multiplexer/demultiplexer combination which uses justification techniques. The Figure shows three inputs being multiplexed together, but any number of inputs could be used. Each input is written steadily into its own elastic store in the form of Fig. 2.

The transmission path bit rate is determined by the clock source shown at top right of the multiplexer. It feeds the frame-length counter which divides the transmitted output into frames and controls a data selector to determine what information will be fitted into each part of the frame. This is laid down by the frame structure (see Section 5). Each frame will contain a frame-alignment word (FAW) which therefore constitutes one input to the data selector. Each input is allocated time slots in the frame; when an input is selected, a gate is opened supplying transmission rate clock pulses to the read clock of that input's elastic store.

If no dummy information is ever sent, the elastic store will gradually empty. A comparator compares the states of the write and read counters. At some point shortly after the beginning of a frame the output of the comparator is examined to determine whether justification — sending a dummy bit — is necessary in that frame. The result of this decision is conveyed in the justification control signal (JCS). When the justifiable time slot (JTS) for the input is reached the output of the corresponding elastic store is selected. If a data bit is sent, the read counter is advanced normally; but if justification is required the read counter is not advanced and a dummy bit is sent. In this latter case, the contents of the store are not decreased in this time slot. To continue with the gas-holder analogy we may suppose that the contents of the gas-holder, measured daily at a particular time, would gradually fall as winter approached. Eventually a point would be reached where gas production would have to be increased to restore the former level.

The processes we have described are reversed in the demultiplexer. There is a frame length counter which must become synchronised with the received frame structure by making use of the frame-alignment words. This being done, the function of each piece of received information is known, and data originally coming from, say, input 1 is directed to the elastic store for output 1. The justification control signal in every frame is identified and decoded so that it is known whether the JTS pertaining to each input contains real or dummy information. All real information is written into the corresponding elastic store, the dummy information being discarded.

The elastic stores in the demultiplexer have an exactly opposite function from those in the multiplexer; they are provided with information in irregular bursts and must feed it out again smoothly. To do this it is necessary to provide the store with a stable, regular read clock which has the same rate as the original input clock which fed the multiplexer. This smooth clock is derived using the fact that information is fed into the demultiplexer elastic store at a rate whose *average* equals the original multiplexer input clock rate. (This must be so, or the multiplexer elastic store would overflow or underflow.) A phase-locked-loop arrangement is used which includes a voltage-controlled

* It is interesting to note that the UK PO have specified positive justification for their digital hierarchy.[10]

oscillator (v.c.o.) and the write and read counters. A comparator detects whether the elastic store is more or less than half full. Its output controls the frequency of the voltage-controlled oscillator via a filter so that the store remains, on the average, half full. The loop bandwidth is made small so that the irregularities at the input are not followed by the v.c.o. from which the read clock is derived.

The fundamental limitation of the system is that it is not possible to smooth out the variations completely; in other words, there will be residual jitter on the output clock. We may think of two types of jitter introduced by the multiplexer/demultiplexer: (i) that arising from the regular gaps in the frame structure (e.g. for the FAW or data from other inputs) and (ii) that arising from justification.

The jitter of type (i) has a fundamental component at frame rate and may therefore be removed by making the loop bandwidth in the demultiplexer somewhat smaller. This is normally quite feasible as the frame rate is usually fairly high to avoid excessively long frames.

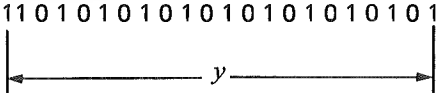
Type (ii) jitter is more of a problem since it may have components at very low frequencies which will not be attenuated by the demultiplexer phase-locked loop. The following example illustrates how these low frequencies arise. Suppose that input and output rates are such as to give a justification ratio of exactly ½. If we write 0 and 1 to denote frames containing real and dummy information in the JTS we obtain the following sequence:

.....0101010101010101010101....

and so on. The resulting jitter will have its lowest frequency component at ½ frame rate — again not very troublesome.

If the ratio of input and output rates is changed very slightly the sequence might, for example, become:

....0101101010101010101010101101010101.....



The sequence now repeats after y frames, in other words the fundamental jitter component is at frame rate/y. As the justification ratio approaches ½ more closely the fundamental frequency becomes lower and lower without limit.

A similar 'beating' effect will occur as the justification ratio approaches any exact fraction. This effect has received considerable attention in the literature.[10,11] It appears to be wise to avoid justification ratios which are near unity, zero or a ratio of integers (a/b) where b is small. Ref. 10 suggests 0·42 (close to 3/7) as a good compromise. In fact we will see in Section 5.2 that the present design problem has insufficient variables to allow a free choice of justification ratio.

5. Design of frame format and framing strategy

5.1. Structure — parallel or serial?

The apparatus has to generate a serial output and is presented with parallel digital video data; it must therefore incorporate elements of both parallel and serial operation. We must choose where to draw the dividing line between them.

The Wyner-Ash error corrector operates in parallel: there are six separate error correctors, one for each level of significance of the video sample word. This configuration has been shown[7] to be desirable because:

(a) It reduces the speed of logic operations, making circuitry both easier and cheaper to build.

(b) When this parallel structure is eventually serialised the six single-error-correcting codes are effectively interlaced. The serial stream thereby acquires a measure of burst-error correction.

The multiplexer proper thus receives one serial input (2048 kbit/s sound) and one parallel (56·75 Mbit/s error-correction-coded video). Two approaches are suggested:

System 1

Serialise the video signal and then, using justification techniques, multiplex the resulting serial video signal with the existing serial sound signal. The multiplex thus has a serial frame structure.

System 2

Convert the serial sound signal to a 6-bit parallel form. This parallel signal is multiplexed with the 6-bit video signal, using justification, to produce a 6-bit, 10 M Word/s signal which is finally serialised to produce a 60 Mbit/s serial output. The multiplex thus has a parallel frame structure.

System 1 might be termed the classical and obvious approach, whereas System 2 at first appears needlessly overcomplicated. We will now discuss the points in favour of System 2.

Parallel operation here has the usual advantage of needing a slower clock rate. This greatly eases any design difficulties, and although extra storage will be required compared with serial operation the total cost is reduced by the much lower cost of slow-speed devices.

Framing is simplified with the parallel structure. There are three* aspects to be considered:

(i) multiplex framing,
(ii) word framing of the video signal,
(iii) parity framing of the video error-correction system.

Multiplex framing is the problem of recognising the frames into which the total multiplexed signal is divided.

* The sound multiplex includes its own framing which need not be considered here.

It is necessary so that demultiplexing and removal of dummy information may take place.

Video word framing relates to the problem of re-grouping video bits into sample words, with the bits in each word in the right order from most to least significant.

Parity framing relates to distinguishing data and parity words in the video bit stream.

If a *serial* structure is used, all three of the above operations will have to be performed separately, and will require some form of framing code to be added to the (in this case) serial video stream in order that video words may be retrieved from it. This framing must be independent from the multiplex framing because of the independence of the video and output bit rates.

With a *parallel* structure these two functions (i.e. (i) and (ii) above) may be combined. The whole frame is made of words of equal length to the video words (6 bits), and sound, frame-alignment words, JCS and JTS are all in 6-bit form. Thus if we specify that video words are inserted with, say, the most significant bit first, then we know that the first video bit following a frame-alignment word is a most significant bit (m.s.b.) and the last video bit before the next FAW is a least significant bit (l.s.b.).

The method of parity framing does not significantly affect the choice of structure since the multiplex frame cannot contain a fixed multiple of 16 video words. Error-corrector framing must therefore be achieved independently and we will discuss the methods of doing this in Section 5.5.

Using parallel operation the justifiable time slots for sound and video become 6-bit words. The amplitude of the jitter resulting from justification will thus be increased, and its frequency lowered. This would normally be an undesirable state of affairs since lower frequency jitter is very difficult to remove with the phase-locked-loop in the demultiplexer. However, the specifications of jitter tolerance for both sound[12] and video[13,14,15] in fact allow for an increase of jitter amplitude with decreasing frequency for very low frequencies. Provided careful attention is paid to the bandwidth of the demultiplexer phase-locked loops, parallel operation need not therefore cause insuperable difficulties with timing jitter.

Parallel structure thus has many advantages without any serious disadvantages; it has therefore been adopted. The block diagram of the sending equipment with this structure is shown in Fig. 4.

5.2. Frame length and related parameters

Having opted for a parallel structure, with multiplexing on a word-by-word basis (where 1 word = 6 bits), we pass to the choice of the multiplex frame length, x words, say. In this situation where all the bit rates are already specified we wish to choose a fairly short frame length in order to minimise both circuit complexity and framing lock-up time.

For the reasons outlined in Section 4.4, positive/negative justification has little to offer for the situation of unrelated bit rates. Positive justification has therefore been specified, requiring 1 JTS word per input per frame.

Although one word per frame could provide JCS signalling for both sound and video justification it seems advisable to protect against error bursts by distributing this information through the frame. It was decided therefore to provide three JCS words, each containing both sound and video JCS information, and to use a majority vote system in decoding. Within these words, which must each contain 2 bits of justification control information (i.e. one for video and one for sound), either further redundancy could be provided or the spare bits could be used for auxiliary signalling.

Each frame will therefore consist of the following components:

y frame-alignment words (FAW) where $y \geqslant 1$
3 justification control signalling words (JCS)
video words, including one justifiable time slot for video (JTSV)
sound words, including one justifiable time slot for sound (JTSS).

It was thought that $y = 2$ would be preferable to $y = 1$ in order to give ample ruggedness of the 60 Mbit/s multiplex.

The number of words in each frame devoted to an input of rate f (including its JTS) is given by the expression:

$(xf)/(60 \cdot 10^6)$ rounded up to the next integer.

The difference between the expression and its rounded up value gives the justification ratio. Substituting the values of f for sound and video we can write an expression giving a lower bound for x:

$$Sx + Vx + 3 + y \leqslant x$$

sound words, video words, JCS, FAW

where $S = \dfrac{2 \cdot 048}{60} = 0 \cdot 034133333$

and $V = \dfrac{56 \cdot 75032}{60} = 0 \cdot 945838667$

i.e. $x \geqslant \dfrac{3 + y}{0 \cdot 020028}$

if $y = 1$ $x \geqslant 119 \cdot 72$

if $y = 2$ $x \geqslant 249 \cdot 65$

These are lower bounds, all higher value integers are not necessarily usable. For example trying $x = 250$:

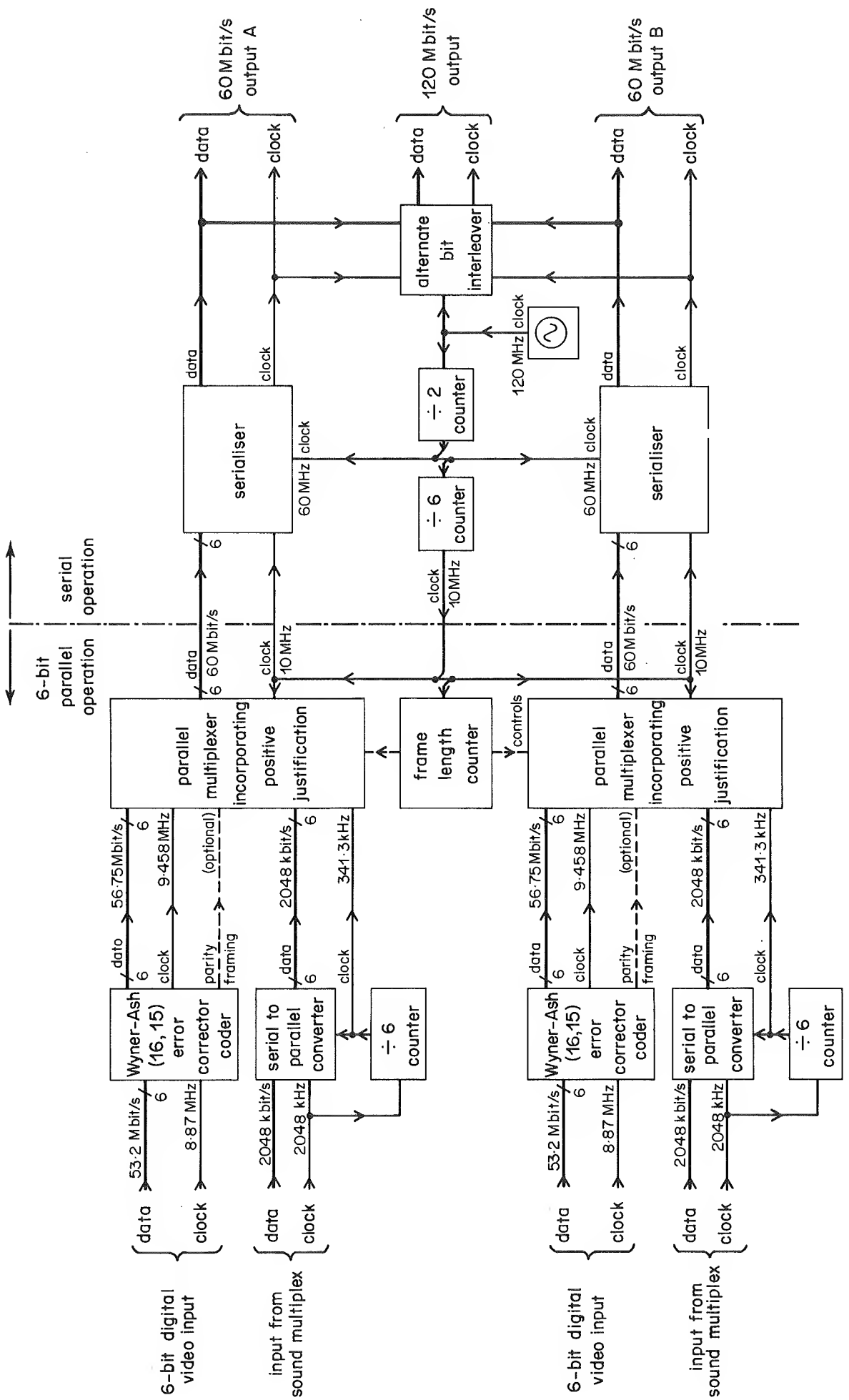


Fig. 4 - Detailed block diagram of sending equipment for 120 Mbit/s operation

S x 250 = 8·533 rounding up to 9 with justification ratio 0·4667
V x 250 = 236·459 rounding up to 237 with justification ratio 0·5403
Total video plus sound words = 246 per frame

Thus there is only room for 3 JCS + 1 FAW.

Working on in this way, restricting the choice to those values of x which conveniently factorise,* we reach x = 300 for which:

S x 300 = 10·24 rounding up to 11 with justification ratio 0·76
V x 300 = 283·7516 rounding up to 284 with justification ratio 0·2484
Total video plus sound words = 295 per frame

leaving room in this case for the desired 3 JCS + 2 FAW. We may note that the nominal justification ratios are close to ¾ and ¼ respectively. These are not ideal, but at least we avoid the worst figure of ½. We do not have control of enough variables (e.g. choosing our own output rate) to optimise the justification ratio as is possible when designing a hierarchical system *ab initio*.

A further check shows that the frequency tolerances of input and output may be safely accommodated.

The next Section describes the detailed structure chosen; we will see that this frame length allows a tidy solution.

5.3. Detailed frame structure

Summarising the preceding Section, we have the following words to locate in the 300-word frame:

10 sound + 1 JTSS
283 video + 1 JTSV
3 JCS
2 FAW

The 2 frame-alignment words may conveniently be located together, effectively forming a bunched 12-bit framing pattern. These are designated words 0 and 1 in the frame (which then runs from word 0 to 299).

The 10 'permanent' sound words may be equally spaced throughout the frame in every 30th word location. This has the advantage that less elastic storage is required than if the sound words were grouped together or spaced irregularly. The remaining 'special' words, i.e. 3 JCS, JTSS, JTSV and 2 FAW may now be placed in the mid-points between sound words. In fact we need only place 'special' words in every other space. Note that the JCS words should precede the JTSS and JTSV. All the remaining word positions are of course allocated to video words. Fig. 5 shows the structure chosen in detail; it is also summarised in the following table:

Word location (from 0 to 299)	Contents	
0	1st frame-alignment word	FAW1
1	2nd frame-alignment word	FAW2
30, 90, 150	Justification control word	JCS
15, 45, 75, 105, 135 165, 195, 225, 255, 285	Sound words	S
210	Justifiable time slot, video	JTSV
270	Justifiable time slot, sound	JTSS
the remaining 283 words	Video words	V

This structure is convenient in that the word locations are easily decoded when the frame length is defined by a ÷300 counter comprising ÷15 and ÷20 counters in cascade.

5.4. Demultiplexing framing strategy

There are two aspects of demultiplexer framing strategy in this particular system.

The basic problem is that the demultiplexer must recognise the frame structure using the frame-alignment words so that it can demultiplex the signal into its component parts. Furthermore, it must do this reliably and quickly even in the presence of transmission errors.

The secondary problem arises when the two 60 Mbit/s streams have been interleaved to produce the 120 Mbit/s

* It is convenient to choose a value of x which factorises in such a way that it is relatively easy to generate the logic waveforms which define the components of the multiplex frame. Moreover, a multiple of 30 is attractive since the ratio S of sound words to the total output is close to 1/30.

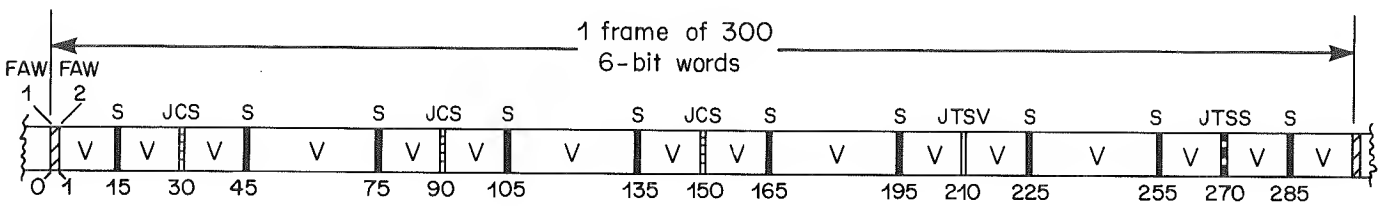


Fig. 5 - Frame structure of the 60 Mbit/s multiplex

signal. It is then necessary to separate the combined signal back into the two streams, say A and B. We must therefore provide means to identify these two streams.

The solution to this problem of channel identification and separation is simple and will be discussed next, as it has a bearing on the detailed strategy of frame-alignment.

The two 60 Mbit/s streams must be synchronous in order that they can be interleaved to produce a 120 Mbit/s output. Considerable simplifications occur if the two channels of the multiplexer share not only a common clock but also common housekeeping waveforms related to the frame structure. Thus, for example, the frame-alignment words occur in the two streams at the same time, and so on for the other time-slots through the frame. If we specify that corresponding bits from the A and B streams are interleaved always placing the bit from the A stream first, the structure is unambiguous. Indeed, it is not even necessary to have different frame-alignment words for the A and B channels.

The general method of framing is simple. Initially, the received stream is searched until the frame code is found, after which it is to be expected that the frame code will recur at regular intervals of one frame length. The signal is examined at these points to confirm the presence of the frame code where it is expected; if the code is consistently absent, then a fresh search must be made.

Points concerning framing which need consideration are:

(i) It is possible for a sequence identical to the frame code to occur in the data — the problem of data 'imitating' the frame code.

(ii) In the presence of transmission errors the correct frame code will not always be present in the expected position although frame synchronisation remains correct. The circuitry must therefore distinguish between such occurrences and genuine loss of synchronisation.

A conventional way of dealing with (ii) above is to count the number of successive framing codes which are incorrect. When a certain value is exceeded synchronisation is deemed to be lost. As this threshold value is increased so the ruggedness of the system is increased.* However, this can lead to a conflict with (i) above, since the time to recognise a false synchronisation produced by imitation is also increased.

A slight modification has been made to this technique whereby effects (i) and (ii) may be catered for independently. Before describing this solution, some further important points particular to our design problem will be listed:

(a) allowance is to be made for 1 x 60 Mbit/s, 2 x 60 Mbit/s and 120 Mbit/s interleaved operation;

(b) circuit design at 120 MHz is limited by the speed of available devices. Furthermore, it is desirable to limit the extent of circuitry operating at 120 MHz and even 60 MHz on the grounds of cost.

Point (a) is dealt with by basing the framing operation at the 60 Mbit/s level. If operation is desired with a 120 Mbit/s interleaved signal then a simple circuit to split this into two 60 Mbit/s streams is used. This 'de-interleaving' operation also helps with point (b) since it requires the minimum of high-speed logic.

Noting that the frame code is effectively 12 bits long (1 x 60 Mbit/s) or 24 bits (2 x 60 Mbit/s), we can afford to make a further simplification in the interests of point (b). It is preferable to search in the serial stream for the frame code, the alternative — operating in the parallel stream — being extremely complicated. However, serial-stream searching would require a long register and much associated fast circuitry. The compromise adopted is to search for only the first frame-alignment word (FAW 1), i.e. one half of the frame code, in the serial stream, and thereby gain synchronisation. Subsequently the presence of the whole frame code where expected is checked, in the serial stream for FAW1 and in the parallel 10 Mword/s stream for FAW2.

The detail of the framing strategy will now be described, assuming in the first instance that one 60 Mbit/s channel is in use. Once FAW1 has been located the presence or absence of FAW1 and FAW2 where they are subsequently expected controls an up/down counter. Its state, say N, is increased by 1 for every incorrect FAW. It is decreased by 1, provided N is not already zero, for every correct FAW. If no errors are present and correct synchronisation has been obtained the counter will quickly clock down to zero and retain that value. If transmission errors are present but synchronisation is maintained then the counter will sometimes clock up, but will quickly clock down to zero. If synchronisation is lost, the counter will steadily clock up and when a threshold value, say T, is reached, synchronisation is deemed to be lost. The process of searching for FAW1 now takes place. When it is found the up/down counter is set to a value, say R, the residue, whence it will eventually clock down to zero if correct synchronisation has been obtained.

Fig. 6 illustrates the process in the form of a flow chart.

The important point to note is the use of the value R; it is the key to separating ruggedness and lock-up time since ruggedness depends on T and lock-up time on $(T-R)$. As T is increased, the transmission error rate required to break synchronisation increases. On the other hand the lock-up time depends on how quickly a false lock caused by imitation of FAW1 is detected. This is quickened by reducing $(T-R)$, the limit being set by the need to avoid loss of a newly-acquired correct lock in the presence of errors. If we increase T but keep $(T-R)$ unaltered, the resistance to transmission errors when in the locked condition increases without any penalty in lock-up time.

* We assume in this discussion that the most likely cause of frame synchronisation loss is the occurrence of random single- or burst-errors in the frame code. This will not always be the case; for example, in digital tape recording the loss of bit synchronisation may be of greater importance and a different strategy is therefore needed.

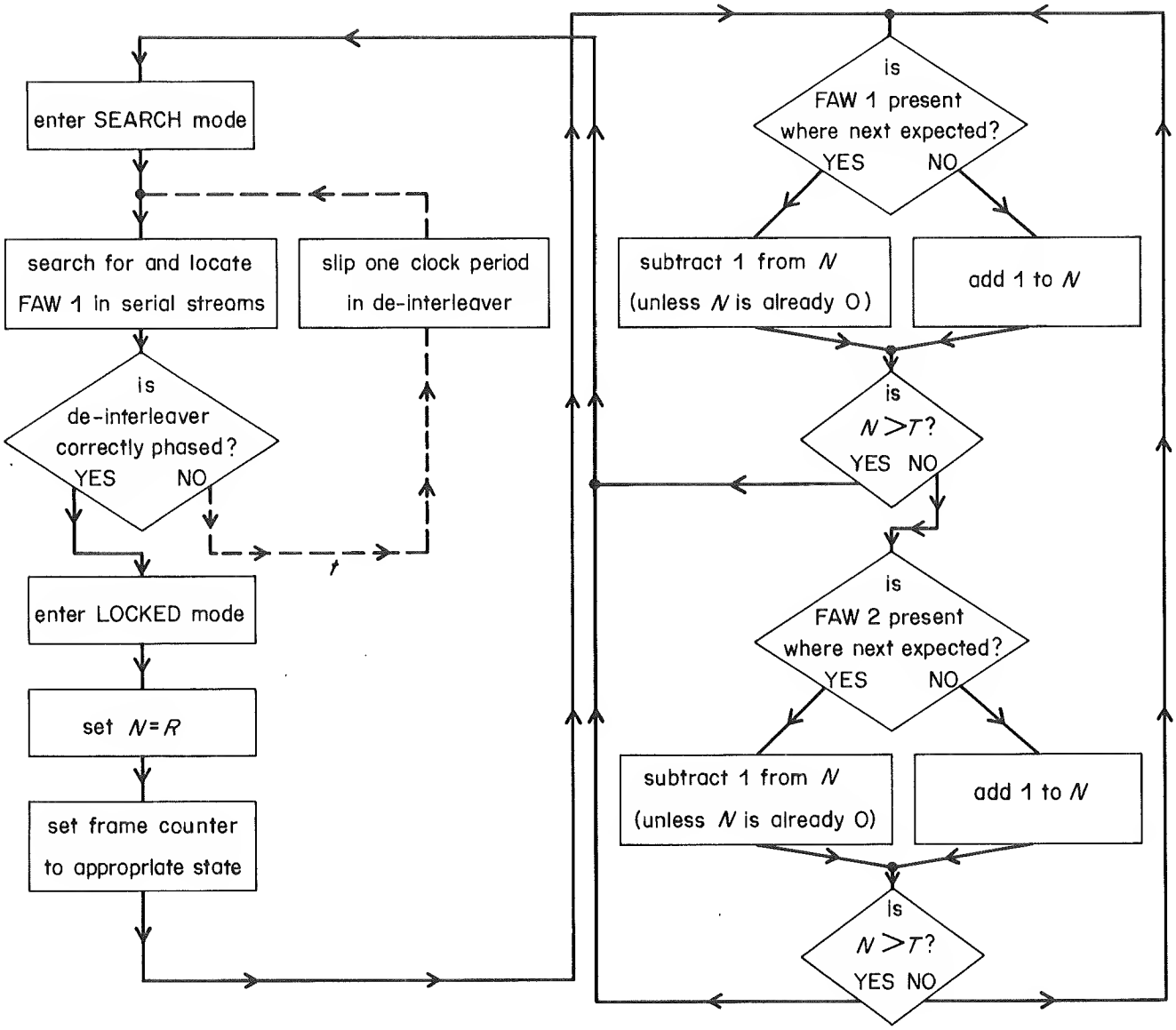


Fig. 6 - Framing strategy flow chart

N – state of up-down counter T – threshold R – residue + only used for interleaved operation

If interleaved operation is in use an extra operation is introduced which is shown in the flow chart as a dotted line. As the de-interleaver arbitrarily splits the incoming stream in two there is no way of knowing which stream is which initially. The search for FAW1 is therefore conducted looking for either correct operation or the 'channels reversed' condition. If the latter is found the de-interleaving operation is interrupted for one clock pulse so that the streams are now sent to their correct destinations. Clearly, this operation must be inhibited once lock is obtained. Note that it is possible to detect the channels reversed condition even if the A and B streams have identical frame-alignment words. This is a result of the time sequence of the bits wherein A always leads B.

5.5. Parity framing of the video error corrector

The error corrector operates in parallel, each of the six bits in each video sample having its own (16, 15) Wyner-Ash coder and decoder. Each of the six Wyner-Ash coders thus located at the sending end regularly produces one parity bit for every 15 data bits fed into it. This is accommodated by having an output clock rate which is (16/15) times the input clock rate, with the parity bits incorporated in the output stream at every 16th bit position. This process takes place in all six error-corrector coders, and a simplification results if all these coders work in step, sharing common control waveforms. Each will then produce a parity bit at the same time so that the parallel output consists of 6-bit words, at (16/15) times the input rate, of which every 16th 6-bit word is now a parity *word*.

The decoder must identify the parity words before correction may proceed. Two general techniques are possible:

(i) to slave the decoder to the coder by an extra signalling channel;

(ii) to deduce the location of the parity words from the information contained in the data and parity bits themselves. The option of using either technique is provided in the apparatus which has been constructed.

The slaving technique, method (i) above, is provided by making use of some of the spare capacity within the JCS words. When this technique is used the JCS redundancy is reduced so that only two of the six bits in each JCS word are used for justification signalling and the remaining four bits are available for other purposes. One of these bits is used to signal parity framing, leaving three bits per JCS word for other auxiliary signalling.

In order to explain the alternative 'self-seeking' technique (method (ii) above) we will first consider the situation when the input to the error corrector decoder equals the coder output, in other words no errors have been introduced in transmission. If the decoder correctly identifies the parity word it will decode the output data correctly, with all the parity checks it performs being satisfied. If it does not correctly identify the parity word, that is to say it performs parity checks over the wrong groups of digits, then these parity checks will frequently not be satisfied. The decoder will then conclude that it has located some errors and will attempt to correct them by inverting the appropriate bits. Excepting the case of certain rare data signals which are related to the period of parity formation (15 data bits), the apparent rate of error correction will be high, approaching the maximum of which the decoder is capable (once per successive 'constraint length'[7] per decoder).

We may now see how to apply this to method (ii) for parity framing. The decoder incorporates circuitry which monitors the apparent rate at which error correction is taking place. When this rate exceeds a pre-set threshold, close to the maximum possible rate, the decoder concludes that it is looking in the wrong place for the parity, slips one word period and tries again. The correct framing phase will be found after a maximum of 15 such slips.

The drawback of this method is that it limits the maximum error-rate which may be corrected. The error-monitoring circuitry mentioned above will of course register when transmission errors are being corrected. If the error-rate is high the threshold may be exceeded although the decoder is correctly framed. The setting of the threshold is thus a compromise: too low a setting will restrict the performance unduly, whereas too high a setting may result in the decoder failing to frame up when the data being conveyed is insufficiently random. Furthermore, for a given data signal, the parity framing lock-up time will increase as the threshold is increased.

In practice method (ii) has proved entirely satisfactory — unwanted framing slips do not occur until the error-rate is so high that the effects on picture quality, despite the use of error correction, are subjectively unacceptable.

When the 'self-seeking' technique is in use, the JCS words may carry justification signalling at maximum redundancy and hence maximum ruggedness. This combination of options has normally been used.

6. Some details of the equipment constructed

6.1. Justification parameters

With the frame format described in Section 5.3, the following points may be noted:

(a) Justification ratios

Allowing input and output frequencies to vary over the tolerances given in Section 2.2 we obtain the following values:

	Justification Ratio		
	min.	nominal	max.
sound	0·7593	0·76	0·7607
video	0·2438	0·2484	0·2530

(b) Storage provided

The elastic stores must accommodate the gaps in the frame structure, likely jitter tolerances and such timing margins as are desirable for good engineering practice. The sizes chosen were 8 words for video and 4 words* for sound.

(c) Bandwidths of demultiplexer phase-locked loops

These are chosen to be as low as possible consistent with predictable design. Jitter due to justification or transmission path disturbances is thus minimised. The values adopted were:

for sound, $f_{-3\,dB}$ = 10 Hz

and for video, $f_{-3\,dB}$ = 7 Hz.

$f_{-3\ dB}$ is the frequency at which sinusoidal disturbances of clock phase are reduced in amplitude by the factor $\sqrt{2}$ owing to the action of the loop. The form of the transfer response is such that jitter attenuation increases with frequency asymptotic to 12 dB/octave. Thus, for example, the video channel phase-locked loop gives over 40 dB jitter attenuation at 100 Hz.

(d) Jitter performance

As the figures in (c) suggest, a very considerable degree of jitter attenuation is effected by the phase-locked loops and no visible or audible impairments due to jitter have been detected, either in the laboratory or in the field trials which were conducted.[1] This is to be expected, since

* Words in each case contain 6 bits. For video they relate to 6-bit data or parity words but for sound they are arbitrary groups of 6 bits formed in the multiplexer in order to work with signals in parallel form (see Section 5.1).

only very low frequency jitter is relatively unattenuated and at such frequencies the otherwise stringent specification for jitter tolerance[12,13,14,15] is relaxed.

6.2. Demultiplexer framing performance

The demultiplexer uses the framing strategy described in Section 5.4 with the facility to vary the threshold T from 1 to 8 and the residue R from 0 to 7.

A brief investigation was made to confirm the general argument of Section 5.4, giving these results:

(i) Observations made over a moderate period of time in which T was varied over the range 2 to 8 showed that the safe error rate,* i.e. that random error rate which just did not cause loss of synchronisation, varied from 4 in 10^5 to > 2 in 10^2.

(ii) With T set at 8, varying R from 0 to 7 in the absence of errors varied the highest observed lock-up times from 400 μs to 200 μs.

(iii) Repeating (ii) with random errors added at a rate of 2 in 10^2 the highest observed lock-up times varied from 450 μs to 250 μs.

These lock-up times are the times taken for the framing circuitry to achieve synchronisation, i.e. for the frame counter to be matched up to the received signal. The total lock-up time before the output clocks are completely stable is longer because the phase-locked loops have to settle down. For video the total settling time (including error-corrector framing, see below) is about 0·3 s, but the output is error-free after about 0·15 s.

The above-mentioned times relate to lock-up following disconnection for a long period. If for some reason synchronisation is lost for only a short period, say, significantly less than 0·1 s, then the phase-locked loop settling time will be correspondingly reduced.

6.3. Error corrector 'self-seeking' framing performance

The error rate threshold involved in this method as described in Section 5.5 can be varied so that the acceptable error rate varies from 3 in 10^4 to 5 in 10^3. With this variation, the measured error-corrector framing time varied from 5 to 12 ms with average picture detail and d.p.c.m. coding; with a p.c.m.-coded sync, burst and black level signal (i.e. an inactive signal) the framing time varied from 7·5 ms to 17·5 ms.

Since even the worst figure (17·5 ms) is fast compared with the demultiplexer phase-locked-loop settling time (0·3 s), there is no real penalty for operating at the most rugged threshold setting (5 in 10^3).

* It is of interest to note that the calculated mean time between synchronisation loss (with T = 8 and 120 Mbit/s operation) for a *random* error rate of 1 in 10^3 is of the order of 1000 years.

6.4. Error monitoring facilities

A method whereby the transmission path bit-error-rate (BER) might be measured was considered to be desirable. In fact most of the necessary circuitry is already present: the methods used for multiplex framing and parity framing (see Sections 5.4 and 5.5) incorporate monitoring of transmission path error performance. The demultiplexer detects erroneous frame-alignment words and the video error-corrector decoder notes every error correction it performs. Thus it is only necessary to add what is in effect a frequency meter with the appropriate choice of timebase to get a direct reading of error rate.

For the 1975 field trials an outlet of error-corrector monitor pulses for connection to an external frequency meter was provided. This was found to be somewhat inconvenient in use and a self-contained built-in unit was subsequently made. This monitors either frame-alignment words or the video error-corrector performance and displays the result in the form

$$j \text{ in } 10^k, \text{ where } j \text{ and } k \text{ are integers.}$$

It can be shown that the accuracy is better than 10% for all usable error rates with either type of monitoring *provided that the errors occur randomly.* Monitoring by means of the error corrector will continue to be accurate for *short** random error bursts, but under these conditions frame-alignment word monitoring will underestimate the true result. Frame-alignment word monitoring must also be treated with circumspection if errors in the transmission path are pattern-dependent since the particular frame-alignment sequence used may be more or less susceptible to errors than the average.

6.5. Extent of built-in flexibility

Some aspects of the built-in flexibility of the apparatus constructed have already been mentioned: error-corrector framing may be directly-signalled or self-seeking, and both multiplex framing and error corrector self-seeking parity framing have adjustable thresholds. The following further options have also been incorporated.

Fig. 4 shows the sending end arrangement as designed to satisfy the original requirements. It can send two complete packages in 120 Mbit/s or alternatively only one in 60 Mbit/s, in each case with one pair of clock and data outlets. One possible re-arrangement, anticipated for satellite transmission experiments,** provides for one 60 Mbit/s package to be connected to a four-phase differential phase-shift-keying modulator which re-

* 'Short' here means 'of burst length less than or equal to the number of error-correcting codes which are interleaved', i.e. 6 bits for the 60 Mbit/s package used singly or 12 bits for the 120 Mbit/s system.

** Such experiments were done with the Post Office over an INTELSAT satellite at Goonhilly Downs in April/May 1976. A full Report is being prepared, but concise descriptions have been published; see Reference 16, for example.

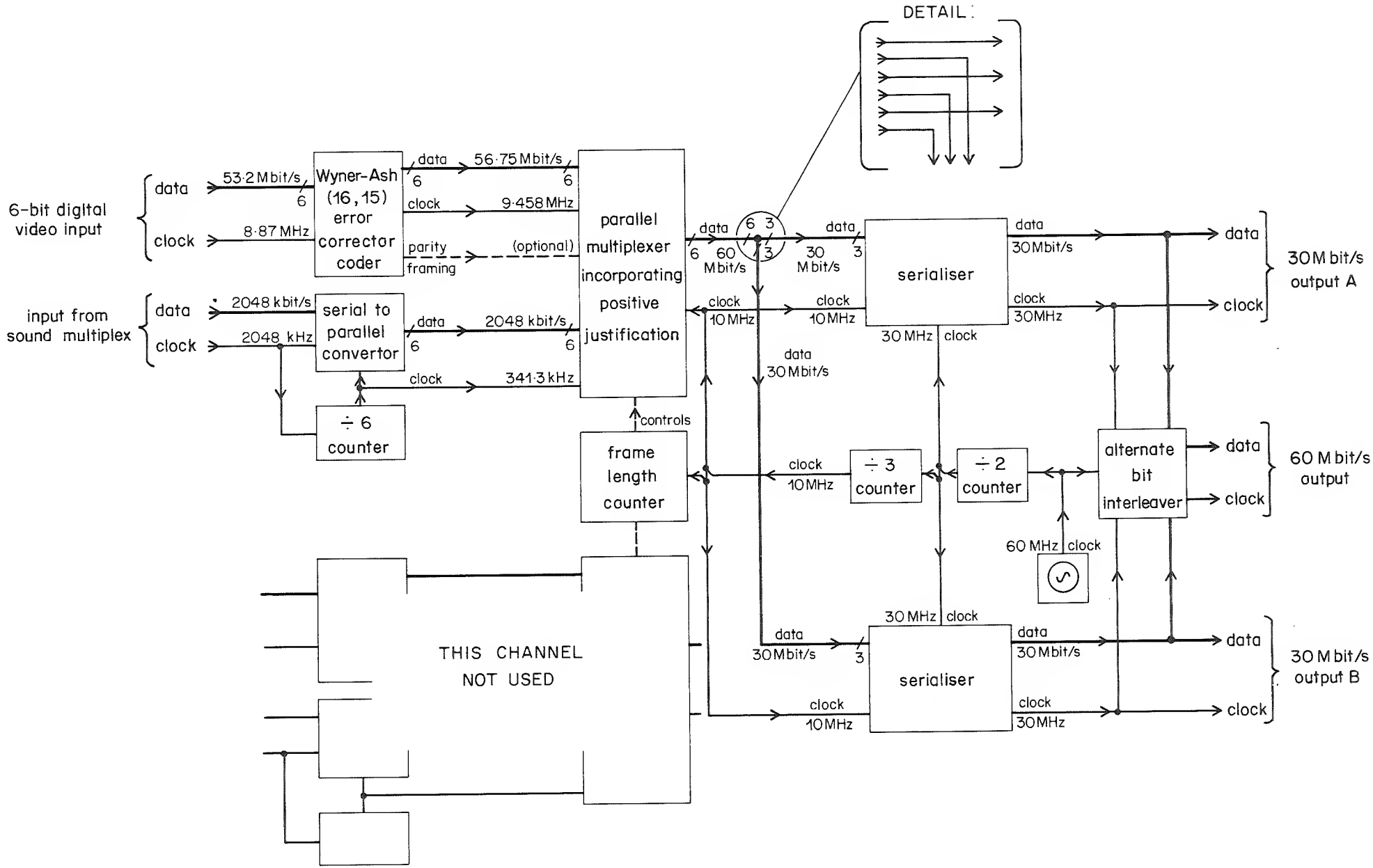


Fig. 7 - Detailed block diagram of sending equipment for 2 x 30 Mbit/s operation

quires two feeds at 30 Mbit/s. This is achieved by the arrangement shown in Fig. 7. The part of the circuitry operating with parallel signals is unchanged, except that, of course, only one channel is required. The connections between the parallel and serial parts of the circuit have been slightly re-arranged so that each serialiser takes in 3-bit words at 10 Mword/s to produce a serial output at 30 Mbit/s. A single 60 Mbit/s signal is also available at the output of the interleaver. It is identical in form to the '60 Mbit/s output A' provided with the original system, (see Fig. 4). The required change of connections can be performed in the experimental apparatus by exchanging one linking plug for another. Other changes are that the master oscillator is now at 60 MHz instead of 120 MHz and the counter which produces parallel word-rate clocks at 10 MHz now divides by 3 instead of 6. In fact, this counter may divide by any chosen factor from 3 to 8 in order to accommodate various word lengths.

Further flexibility is provided in the parallel circuitry. Although we have always discussed 6-bit operation so far, the construction is such that word lengths up to 8 may be used. The factor by which a counter divides the 2048 kbit/s sound clock may also be varied from the normal value of 6 to suit other word lengths.

We have discussed possible changes to the sending equipment; the corresponding flexibility is of course available in the receiving equipment.

7. Conclusions

It has been found to be practicable to build up a 60 Mbit/s digital transmission package which carries one System I PAL colour video signal and six high-quality sound signals, with transmission error-protection for the video and the audio signals. Two of these 60 Mbit/s packages have been combined for transmission at 120 Mbit/s.

The bit rates produced in coding these video and sound signals were independent of each other and of the transmission rate. In consequence a technique called justification was required when the signals were multiplexed. The design of equipment which performs this multiplexing operation has been described.

The equipment constructed was used successfully in field trials in 1975 with an experimental Post Office digital line transmission system. The design included an element of flexibility so that field trials with other transmission systems, e.g. the 60 Mbit/s satellite experiments done in April/May 1976, could use the same equipment.

8. References

1. OSBORNE, D.W. 1976. Experimental digital transmission of multiplexed video and audio signals at 120 Mbit/s. BBC Research Department Report No. 1976/16. An edited version of this Report is a paper at the International Broadcasting Convention 1976; IEE Conference Publication No. 145, pp. 84 – 88.

2. DEVEREUX, V.G. *and* MEARES, D.J. 1972. Pulse code modulation of video signals: subjective effect of random digit errors. BBC Research Department Report No. 1972/14.

3. MEARES, D.J. 1971. Pulse code modulation of video signals: parallel-to-serial and serial-to-parallel conversion. BBC Research Department Report No. 1971/14.

4. DEVEREUX, V.G. *and* PHILLIPS, G.J. 1974. Bit-rate reduction of digital video signals using differential p.c.m. techniques. International Broadcasting Convention 1974. IEE Conference Publication No. 119, pp. 83 – 89.

5. CROLL, M.G., OSBORNE, D.W. *and* SPICER, C.R. 1974. Digital sound signals: the present BBC distribution system and a proposal for bit-rate reduction by digital companding. Ibid., pp. 90 – 95.

6. DEVEREUX, V.G. 1975. Digital video: sub-Nyquist sampling of PAL colour signals. BBC Research Department Report No. 1975/4.

7. STOTT, J.H., OLIPHANT, A. *and* OSBORNE, D.W. 1974. Digital video: error-correcting codes and a practical study of a Wyner-Ash error corrector. BBC Research Department Report No. 1974/40.

8. RATLIFF, P.A. 1974. Digital video: subjective assessment of an experimental Wyner-Ash error corrector. BBC Research Department Report No. 1974/41.

9. MOORE, T.A. 1974. Digital transmission codes: properties of HDB3 and related ternary codes with reference to broadcast signal distribution. *The Radio and Electronic Engineer*, 1974, **44**, 8, pp. 421 – 427.

10. LEVETT, F.A.W., *and* STURCH, P.M.E. 1975. Principles of asynchronous digital multiplexing. IEE Conference on Telecommunications Transmission 1975, IEE Conference Publication No. 131, pp. 79 – 81.

11. DUTTWEILER, D.L. 1972. Waiting time jitter. *Bell System Technical Journal*, **51**, 1, pp. 165 – 207.

12. MANSON, W.I. 1974. Digital sound signals: subjective effect of timing jitter. BBC Research Department Report No. 1974/11.

13. DEVEREUX, V.G. 1971. Pulse code modulation of video signals: subjective tests on acceptable limits for timing jitter in the decoded analogue samples. BBC Research Department Report No. 1971/42.

14. DEVEREUX, V.G. *and* WILKINSON, G.C. 1973. Digital video: effect of PAL decoder alignment on the acceptable limits for timing jitter. BBC Research Department Report No. 1973/1.

15. CHEW, J.R. *and* STEBBINGS, D.W. 1975. Specification of the timing stabilities of broadcast monochrome and PAL colour television signals. BBC Research Department Report No. 1975/14.

16. Digital colour TV by satellite. *Wireless World.* 1976, **82**, 1488, p. 71.

Printed by BBC RESEARCH DEPARTMENT, Kingswood Warren, Tadworth, Surrey, KT20 6NP